REMARKS

Claims 1-21 are pending in the application.

Claims 1, 2, 4-16, and 18-21 are currently amended. Applicants respectfully submit that no new matter is added to currently amended claims 1, 2, 4-16, and 18-21.

Claims 8 and 10-15 stand rejected under 35 U.S.C. §101.

Claims 1-21 stand rejected under 35 U.S.C. §103(a) as unpatentable over "A Tree-Based Statistical Language Model for Natural Language Speech Recognition" by Bahl et al., hereinafter, Bahl, in view of U.S. Patent No. 6,292,772 to Kantrowitz.

Applicants respectfully traverse the rejections based on the following discussion.

I. The 35 U.S.C. §101 Rejection

Claims 8 and 10-15 stand rejected under 35 U.S.C. §101 because the Office Action asserts that the claimed invention is directed to non-statutory subject matter, i.e., "a computer program product".

Independent claim 8 is currently amended to recite in relevant part, "A program storage device readable by machine, tangibly embodying a program of instructions executable by said machine to perform a method for language modelling of mixed language expressions, said method comprising: ... ." in accordance with the ruling of In re Beauregard, 53 F.3d 1583 (Fed. Cir. 1995). Dependent claims 10-15 are currently amended to provide proper antecedent basis to currently amended, independent claim 8.

For at least the reasons outlined above, Applicants respectfully submit that claims 8 and 1-15, as currently amended, satisfy the statutory requirements of 35 U.S.C. §101. Withdrawal of the rejection of claims 8 and 10-15 under 35 U.S.C. §101 is respectfully solicited.

II. The 35 U.S.C. 103(a) Rejection over Bahl and Kantrowitz

A. The Bahl Disclosure

Bahl discloses that in any practical natural-language system with even a moderate vocabulary size, it is clear that the language model probabilities Pr $\{ w_i \mid w_1, w_2, \ldots, w_{i-1} \}$ cannot

be stored for each possible sequence $w_1, w_2, \ldots, w_{i-1}$. Even if the sequences were limited to one or two sentences in length, the number of distinct sequences would be so large that a complete set of probabilities could not be computed, never mind stored or retrieved. To be practicable, then, a language model must have many fewer parameters than the total number of possible sequences $w_1, w_2, \ldots, w_{i-1}$. An obvious way to limit the number of parameters is to partition the various possible word histories $w_1, w_2, \ldots, w_{i-1}$ into a manageable number of equivalence classes. (Page 1001, col. 2, which is cited by the Office Action).

A simple-minded, but surprisingly effective, definition of equivalence classes can be found in the N-gram language models [5], [10]. In this model, word sequences are treated as equivalent if and only if they end with the same N-1 words. Typically N=3, in which case the model referred to as a 3-gram or a trigram model. The trigram model is based upon the approximation

$$\Pr\{ w_i \mid w_1, w_2, \ldots, w_{i-1}\} \approx \{ w_i \mid w_1, w_{i-2}, w_{i-1}\}, \qquad (4)$$

which is clearly inexact, but apparently quite useful. Maximum-likelihood estimates of N-gram probabilities can be obtained from their relative frequencies in a large body of training text. But since many legitimate N-grams are likely to be missing from the training text, it is necessary to "smooth" the maximum-likelihood estimates so as to avoid probabilities of zero. The trigram model can be smoothed in a natural way using the bigram and unigram relative frequencies as described in [6]. (Page 1001, col. 2, which is cited by the Office Action).

Bahl also discloses a tree-growing algorithm in which the set $S^c{}_i$ minimizes the average conditional entropy at the current node. Determining $S^c{}_i$ amounts to partitioning the values taken by X_i [a discrete random variable] into two groups: those in $S^c{}_i$ and those not in $S^c{}_i$. There is no known practical way of achieving a certifiably optimal partition, especially in applications like language modeling where X_i can take a large number of different values. As before, the best realistic hope is to find a "good" set $S^c{}_i$ via some kind of heuristic search. Possible strategies range from relatively simple greedy algorithms to the computationally expensive techniques of simulated annealing. (Page 1006, first paragraph, which is cited by the Office Action).

Let X denote the set of values taken by the variable X. In our case, X is the entire vocabulary. The following algorithm determines a set S in a greedy fashion.

1) Let S be empty.

2) Insert into S the $x \varepsilon X$ which leads to the greatest reduction in the average conditional entropy (7). If no $x \varepsilon X$ leads to a reduction, make no insertion.

3) Delete from S any member x, if so doing leads to a reduction in the average conditional entropy.

4) If any insertions or deletions were made to S, return to step 2. (Page 1006, paragraphs 2 and 3, which are cited by the Office Action).

B. The Kantrowitz Disclosure

Kantrowitz discloses a method of recognizing the language of a single word as to spelling and grammar corrections (e.g., identifying the appropriate language resources on a document, paragraph, sentence or even individual word basis), the automatic invocation of transliteration software based on the language of the words (e.g., automatic ASCII to Kanji substitution without requiring the user to explicitly switch into a Kanji mode), the automatic invocation of appropriate machine translation tools when the document's language is different from the user's native tongue(s), the use of document language identification to eliminate from database or web search results any documents which are not written in the user's native language and automatic identification of user-appropriate languages for the user interface. (Abstract).

Kantrowitz also discloses that his invention allows a user to type in English or Romanji as needed, with the system automatically distinguishing between the two and converting the Romanjito Kanji as necessary. In a mixed-language document, this regular expression can be used to select the appropriate dictionary and thesaurus for use with the word. ... The invention is able to identify the language of individual words in isolation with high accuracy. The accuracy in identifying the language of individual words typically is equal to that of whole-document language identification systems. Moreover, the ability to identify the language of individual words permits document processing resources to be applied on a word-by-word basis. (col. 6, lines 11-51, which is cited by the Office Action).

C. Arguments

Currently amended, independent claims 1 and 8 recite in relevant part,

"storing word equivalence probabilities relating to words of a first language and words in at least one other language;

generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ;

generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and

determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities".

Similarly, currently amended, independent claim 9 recites in relevant part,

"a memory for storing word equivalence probabilities relating to words of a first language and words in at least one other language; and

a processor configured to:

generate a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ;

generate monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and

determine a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities".

Bahl merely discloses a standard statistical language model without disclosing how to handle mixed-language data.

In addition, nowhere does Bahl disclose, teach or suggest storing word equivalence probabilities. Instead, Bahl describes partitioning word histories into equivalence classes, such as, *N*-grams, without mentioning mixed-language data. Bahl's disclosure relates to a tree-based statistical language model for natural language speech recognition <u>in but one language</u>. (emphasis added).

Nowhere does Bahl disclose, teach or suggest at least the present invention's features of: "storing word equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claims 1 and 8; and "a memory for storing word equivalence probabilities relating to words of a first language and words in at least one other language; and a processor configured to: generate a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generate monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determine a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claim 9.

Instead, Bahl merely discloses a standard statistical language model without disclosing how to handle mixed-language data.

Kantrowitz merely discloses identifying individual words of mixed languages in a document.

Nowhere does Kantrowitz disclose, teach or suggest "storing word equivalence probabilities relating to words of a first language and words in at least one other language", as recited in claims 1 and 8 of the present invention and as similarly recited in claim 9. (emphasis added).

Furthermore, since Kantrowitz does not disclose "word equivalence probabilities" of a first language and at least one other language, Kantrowitz cannot "generat[e] a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities", nor can Kantowitz "determin[e] a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and

the stored word equivalence probabilities", as recited in independent claims 1 and 8 and similarly recited in independent claim 9. (emphasis added).

Nowhere does Kantrowitz disclose, teach or suggest at least the present invention's features of: "storing word equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claims 1 and 8; and "a memory for storing word equivalence probabilities relating to words of a first language and words in at least one other language; and a processor configured to: generate a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generate monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determine a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claim 9.

Instead, Kantrowitz merely discloses identifying individual words of mixed languages in a document.

For at least the reasons outlined above, Applicants respectfully submit that Bahl and Kantrowitz, either individually or in combination, do not disclose, teach or suggest at least the present invention's features of: : "storing word equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claims 1

and 8; and "a memory for storing word equivalence probabilities relating to words of a first language and words in at least one other language; and a processor configured to: generate a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities ... ; generate monolingual next word hypothesis probabilities in the first language based upon the monolingual word history ... ; and determine a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities", as recited in independent claim 9. Accordingly, Bahl and Kantrowitz, either individually or in combination, fail to render obvious the subject matter of independent claims 1, 8, and 9, and dependent claims 2-7 and 10-21 under 35 U.S.C. §103(a). Withdrawal of the rejection of claims 1-21 under 35 U.S.C. §103(a) as unpatentable over Bahl and Kantrowitz is respectfully solicited.

III. Formal Matters and Conclusion

Claims 1-21 are pending in the application.

Applicants respectfully submit that claims 8 and 1-15, as currently amended, satisfy the statutory requirements of 35 U.S.C. §101.

With respect to the rejections of the claims over the cited prior art, Applicants respectfully argue that the present claims are distinguishable over the prior art of record. In view of the foregoing, the Examiner is respectfully requested to reconsider and withdraw the rejections to the claims.

In view of the foregoing, Applicants submit that claims 1-21, all the claims presently pending in the application, are patentably distinct from the prior art of records and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest time possible.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary.

Please charge any deficiencies and credit any overpayments to Attorney's Deposit Account Number 09-0441.

Respectfully submitted,

Dated: July 29, 2008

/Peter A. Balnave/
Peter A. Balnave, Ph.D.
Registration No. 46,199

Gibb & Rahman, LLC
2568-A Riva Road, Suite 304
Annapolis, MD 21401
Voice: (410) 573-5255
Fax: (301) 261-8825
Email: Balnave@Gibb-Rahman.com
Customer Number: 29154